Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 6919–6923, August 1992
Biochemistry

Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates

(fidelity/mutagenesis/reverse transcription/retroviruses)

JAYNE C. BOYER, KATARZYNA BEBENEK, AND THOMAS A. KUNKEL*

Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709

Communicated by Howard M. Temin, April 27, 1992 (received for review January 9, 1992)

ABSTRACT **Sequence variation in the type 1 human immunodeficiency virus (HIV-1) results, in part, from inaccurate replication by reverse transcriptase. Although this enzyme is error-prone during synthesis *in vitro* with DNA templates, the fidelity of RNA-dependent DNA synthesis relevant to minus-strand replication in the virus life cycle has not been examined extensively. In the present study, we have developed a system to determine the fidelity of transcription and reverse transcription and have used it to compare the fidelity of DNA synthesis by the HIV-1 reverse transcriptase with RNA and DNA templates of the same sequence. Overall, fidelity was several-fold higher with RNA than with DNA. Sequence analysis of mutants generated with the two substrates revealed that differences in error rates were substantial for specific errors. Fidelity with RNA was >10-fold higher for substitution and minus-one nucleotide errors at five different homopolymeric positions. Because such errors likely result from template–primer slippage, this result suggests that misaligned intermediates are formed and/or used less frequently with an RNA template–DNA primer than with a DNA template–DNA primer. The results also suggest that HIV-1 reverse transcriptase synthesis with an RNA template–DNA primer was error-prone during incorporation of the first two nucleotides, perhaps due to aberrant enzyme–substrate interactions as synthesis initiates. The unequal error rates with RNA and DNA templates suggest that mistakes during minus- and plus-strand DNA synthesis may not contribute equally to the mutation rate of HIV-1. The data also provide estimates of substitution and frameshift error rates during transcription by T7 RNA polymerase.**

The causative agent of AIDS, human immunodeficiency virus type 1 (HIV-1), exhibits extensive genomic heterogeneity (1). One source of this sequence diversity is inaccurate DNA replication by its reverse transcriptase (RT). Conversion of the viral genomic RNA to double-stranded DNA is a two-step process. The minus strand is synthesized by using viral RNA as a template. The RT then uses this newly synthesized, complementary DNA as a template for second-strand synthesis.

HIV-1 RT, which lacks a 3′ → 5′ exonuclease proofreading activity (2), has been shown to be error-prone during DNA-dependent DNA synthesis *in vitro* (2–7). Although such studies are relevant to second-strand synthesis, the fidelity of reverse transcription of RNA to DNA is also relevant to retroviral mutagenesis. Limited information is available on the fidelity of RNA-dependent DNA synthesis by HIV-1 RT (8–10). We therefore decided to examine HIV-1 RT fidelity with heteropolymeric RNA by adapting a forward mutation assay (2) previously used to establish HIV-1 RT error rates with a DNA template (4). This assay has permitted a direct comparison of fidelity of HIV-1 RT with RNA and DNA templates of the same sequence as well as an estimate for error rates during transcription by T7 RNA polymerase.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Enzymes and Reagents. The recombinant form of HIV-1 RT (p66/p51 heterodimer) has been described (2). Avian myeloblastosis virus (AMV) RT (specific activity was 100,000 units per mg) was from Boehringer Mannheim. T4 polynucleotide kinase and T7 polymerase were from United States Biochemical. Oligonucleotides were from Research Genetics (Huntsville, AL).

Fidelity Assay. The forward mutation assay scores errors in the *lacZα* gene in bacteriophage M13mp2 (2). As originally designed for DNA-dependent synthesis, correct polymerization to fill a 390-nucleotide (nt) gap in an otherwise double-stranded circular molecule produces DNA that yields dark-blue M13 plaques upon transfection of an appropriate *Escherichia coli* host strain. Errors yield lighter blue or colorless plaques, detected at 114 template positions for base substitutions and 150 positions for frameshifts. To examine fidelity with an RNA template, we adapted the assay as described below.

Transcription Reaction Conditions. Transcription reactions (100 μl) contained 40 mM Tris·HCl (pH 7.5), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM dithiothreitol, RNasin ribonuclease inhibitor at 1 unit per μl (Promega), 500 μM ATP, 500 μM GTP, 500 μM CTP, 500 μM UTP, 2–10 μg of *Fsp* I-linearized M13mp2 DNA containing the T7 RNA polymerase promoter, and 38 units of T7 RNA polymerase (Promega). Reactions were incubated for 2 hr at 37°C. The DNA template was digested with RQ1 RNase-free DNase (Promega) (DNA at 1 unit per μg) for 15 min at 37°C. The RNA was purified by phenol extraction and ethanol precipitation, redissolved in RNase-free water, and stored at −70°C.

DNA- and RNA-Templated DNA Synthesis Reactions. DNA-dependent gap-filling synthesis reactions (50 μl) contained 20 mM Tris·HCl (pH 7.8), 10 mM $MgCl_2$, 2 mM dithiothreitol, 1 mM dATP, 1 mM dGTP, 1 mM dTTP, 1 mM dCTP, 67 fmol of gapped DNA molecules (300 ng), and 2 pmol of HIV-1 RT (establishing a 30:1 enzyme/template ratio). Reactions were incubated at 37°C for 1 hr and then analyzed as described (9), confirming that the gap had been filled. cDNA synthesis reactions (50 μl) contained 20 mM Tris·HCl (pH 7.8), 10 mM $MgCl_2$, 2 mM dithiothreitol, 1 mM dATP, 1 mM dGTP, 1 mM dTTP, [α-^{32}P]dCTP at 6000 cpm/pmol, RNase ribonuclease inhibitor at 1 unit/μl (Promega), 1 pmol of RNA molecules (100 ng) primed with a 2-fold molar excess of a 15-mer DNA oligonucleotide, and 30 pmol of HIV-1 RT (for Table 1, experiments 1 and 2, again establishing a 30:1

Abbreviations: HIV-1, human immunodeficiency virus type 1; RT, reverse transcriptase; AMV, avian myeloblastosis virus; nt, nucleotide(s).

*To whom reprint requests should be addressed.

enzyme/template ratio). For experiment 3 in Table 1, 80-fold less HIV-1 RT was used. The RNA was first heated to 65°C for 5 min and cooled on ice before adding to the other reagents. The reaction was incubated at 37°C for 2 hr and stopped by adding EDTA to 15 mM. After denaturation at 80°C for 5 min and cooling on ice, the RNA was digested with RNase A and RNase T1 for 1 hr at 37°C. The cDNA was purified by phenol extraction, precipitated with ethanol, and resuspended in sterile H_2O. An aliquot was analyzed by electrophoresis in a 4% denaturing polyacrylamide gel followed by autoradiography. Sequencing reaction products were used as molecular weight markers. The amount of cDNA produced was estimated by cutting and counting radioactive bands from the gel. The 5′ end of the remaining cDNA was phosphorylated in a 50-μl reaction containing 50 mM Tris·HCl (pH 8.0), 10 mM $MgCl_2$, 1.5 mM spermidine, 1 mM ATP, and 0.5 unit of T4 polynucleotide kinase per pmol of DNA. Reactions were incubated for 1 hr at 37°C and terminated by heating to 65°C for 5 min.

Hybridizing cDNA Fragment to Gapped DNA. A 5-fold molar excess of cDNA was mixed with gapped M13mp2 DNA in 300 mM NaCl/30 mM sodium citrate, heated to 70°C, and slowly cooled to room temperature. The product was analyzed by electrophoresis in an agarose gel as described (11).

RESULTS

We previously determined the fidelity of HIV-1 RT during DNA-dependent gap-filling synthesis by using an assay that scores errors in the *lacZα* gene in bacteriophage M13mp2 (2). To examine fidelity with an RNA template of this same sequence, we adapted the assay as shown in Fig. 1. The promoter for transcription by T7 RNA polymerase was placed adjacent to the *lacZα* gene. After digestion of double-stranded DNA with restriction endonuclease *Fsp* I, run-off transcription produced the desired RNA template. An oligonucleotide DNA primer was used to initiate cDNA synthesis by the RT, producing a cDNA molecule spanning the same *lacZα* target (positions −84 through 174) used for scoring errors with the gapped DNA template. Errors were recovered by hybridizing the cDNA strand to a gapped circular DNA substrate and transfecting this into an *E. coli* α-complementation host.

Run-off transcription by T7 RNA polymerase should generate a transcript of 313 nt (from position −118 through +195). Fig. 2*A* shows that the size of the transcript actually obtained was slightly longer than 300 nt. The product of complete cDNA synthesis by HIV-1 RT should be 292 nt in length (from position −118 through +174). Fig. 2*B* shows that HIV-1 RT synthesized a product of the expected length under reaction conditions similar to those used for DNA-dependent synthesis—i.e., the same enzyme-to-template ratios, pH, $MgCl_2$, and dNTP concentrations. A longer exposure of the autoradiogram revealed a small amount of shorter-length products.

The results of hybridization of this cDNA to a circular M13mp2 substrate containing a 390-nt gap, as analyzed by electrophoresis in an agarose gel, are illustrated in Fig. 2*C*. Lane 1 shows the position for a nicked, fully double-stranded standard DNA (form II), lane 2 shows the position of the DNA substrate with an unfilled 390-nt gap, and lane 3 shows the product of hybridization of a 5-fold molar excess of the cDNA to the gapped substrate. The presence of a single band at the expected position indicated that essentially all gapped DNA has received a cDNA copy.

Rescuing cDNA synthesis errors by hybridization requires expression of the mutant phenotype upon transfection. To examine the efficiency of error expression by this procedure, the cDNA generated from a wild-type—i.e., blue-phenotype—RNA template was hybridized to three different gapped DNAs, containing either a single-base substitution, 1-nt addition, or 1-nt deletion mutation (each having a colorless phenotype). These heteroduplexes contain mispairs or extra nucleotides that mimic the error-containing products generated by cDNA synthesis. The colors are ideal for discrimination on plates, and the symmetry of the misalignment (whether the extra nucleotide is in the minus-strand or the plus-strand) or the mispair [e.g., T(−)·G(+) versus G(−)·T(+)] does not matter for minus-strand expression values (11). The resulting "blue–white" heteroduplexes each yielded ≈50% blue plaques upon transfection. This efficiency was similar to expression values of ≈60% for DNA-dependent DNA synthesis errors (11).

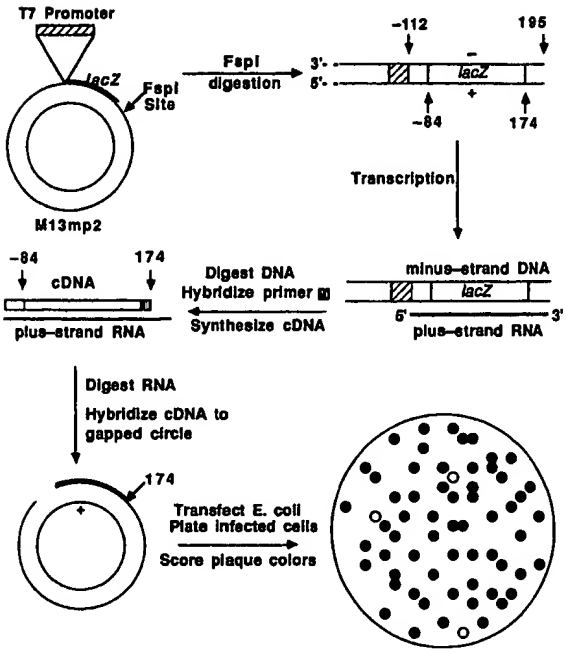

FIG. 1. Protocol for determining fidelity of DNA synthesis with an RNA template. With oligonucleotide-directed mutagenesis (12), a transcription vector containing the promoter for T7 RNA polymerase placed between nt −112 and −113 (where position +1 is the first transcribed nucleotide of the *lacZα* gene) was constructed. After digestion of the double-stranded DNA with restriction endonuclease *Fsp* I, transcription by T7 RNA polymerase produces a 313-nt transcript, the 5′ end of which is nt −118 and the 3′ end is nt 195. By using the RNA as a template, cDNA synthesis was initiated from a 15-mer DNA primer the 5′ end of which is nt 174. After heat denaturation, the RNA was digested, and an excess of cDNA was used for hybridization to a circular DNA substrate containing a single-stranded gap. The resulting molecule contained one phosphodiester bond interruption between nt 174 and 175 and a 121-nt single-stranded gap. This gap contains the nucleotide sequence for the C-terminal end of the *E. coli lacI* gene (the termination codon of which ends at nt −84) and is, therefore, not part of the *lacZα*-complementation target sequence for scoring errors. Neither the nick nor the gap significantly diminished the probability that changes in the *lacZα*-complementation target sequence of the hybridized cDNA were scored as mutants upon transformation and plating (see *Results*).

Transfection of competent cells with the products of DNA-dependent gap-filling reactions by HIV-1 RT yielded *lacZα*-complementation mutant frequencies similar to those obtained in an earlier study (Table 1, DNA template). Transfection with the products of RNA-dependent DNA synthesis reactions yielded several-fold lower mutant frequencies (Table 1). These data demonstrate that the overall fidelity of

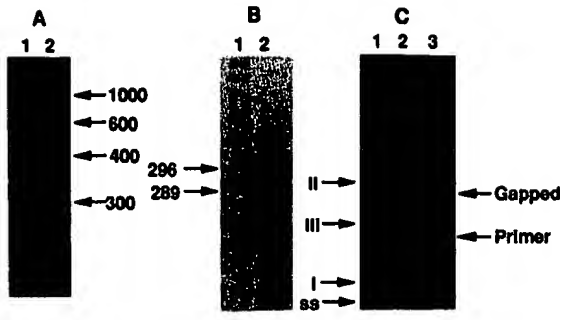

FIG. 2. Electrophoretic analysis of reaction products. (*A*) Product of transcription by T7 RNA polymerase was analyzed by electrophoresis in a denaturing 4% polyacrylamide gel followed by staining with ethidium bromide. Lanes: 1, RNA transcript; 2, 0.3- to 1.6-kilobase (kb) RNA markers (Boehringer Mannheim). (*B*) Products of cDNA synthesis by HIV-1 RT (Table 1, experiment 1) were analyzed by electrophoresis in a denaturing 4% polyacrylamide gel followed by autoradiography. Lanes: 1, sequencing ladder size markers; 2, cDNA. (*C*) Product of hybridization of cDNA to gapped circular M13mp2 DNA was analyzed by electrophoresis in a 0.8% agarose gel containing ethidium bromide at 0.5 μg/ml. Lanes: 1, DNA standards—II (relaxed circular DNA), III (linear double-stranded DNA), I (covalently closed supercoiled DNA), ss (single-stranded circular DNA); 2, gapped DNA before hybridization of cDNA and (residual) 6804-base-pair linear primer DNA used to construct gapped DNA; 3, gapped DNA after hybridization with cDNA, as described.

DNA synthesis with an RNA template was higher than with a DNA template of the same sequence.

A characteristic of error-prone synthesis by HIV-1 RT with a DNA template is the dependence of error rates on local nucleotide sequence. High error rates were observed for single-base substitutions and 1-nt deletions at specific DNA template positions (see figure 1 in ref. 4). To determine whether this was the case with the RNA template, we analyzed the sequences of 139 mutants recovered from the cDNA synthesis reactions. The results include mutants containing only a single-nucleotide change (Fig. 3) as well as mutants (data not shown, but see the legend to Fig. 3) that contained two changes, deletions, or complex changes. This sequence information can be used to determine quantitative error rates per detectable nucleotide polymerized for the positions found to be hot spots with either template (Table 2 and see legend for calculations).

Table 1. Mutation frequencies for reactions with DNA and RNA template

RT	Plaques, no. Total	Plaques, no. Mutant	Mutation frequency × 10^4
DNA template			
HIV-1 experiment 1	7,355	397	540
HIV-1 experiment 2	11,350	386	340
HIV-1 experiment 3			390*
RNA template			
HIV-1 experiment 1	5,667	120	210
HIV-1 experiment 2	12,728	273	210
HIV-1 experiment 3	14,329	130	91
AMV	39,020	217	56

*Average of three independent determinations (4). The mean and SD for all five determinations was 410 ± 87 × 10^{-4}.

This analysis reveals substantial differences in error rates with RNA and DNA templates. To the extent that the mutant frequency for any error is lower with RNA than DNA, reverse transcription is more accurate than DNA-dependent synthesis. Thus, HIV-1 RT was 73-fold more accurate for U → C errors at position −36 in the RNA template than for the corresponding T → C substitution error in the DNA template (Table 2). Likewise, G → A mutations occurred 14-fold less frequently at position 90 with RNA as compared with DNA. Excluding these sites, the overall average substitution frequency for other sites was similar with RNA and DNA. However, HIV-1 RT was more accurate with RNA for 1-nt deletions at four different homopolymeric template sequences and for 1-nt additions at a template UUUU run (Table 2).

There are three possible sources for the mutants that were recovered with the RNA template. Those present in the DNA preparation used for transcription should be rare because the mutant frequency of the starting DNA, 3 × 10^{-4}, was much lower than the mutant frequencies in Table 1. In an attempt

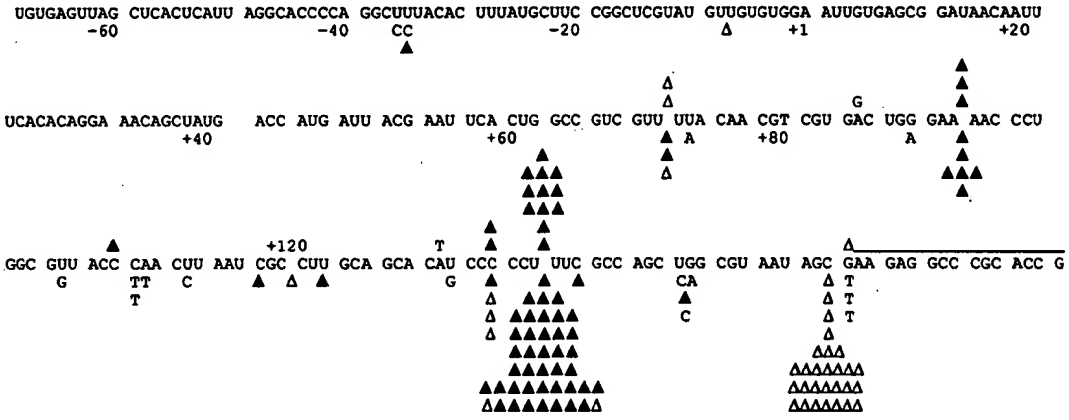

FIG. 3. Single-base mutations by HIV-1 and AMV RT with an RNA template. The primary DNA sequence of the viral (+)-template strand from positions −69 to 174 is shown. Mutants shown above each line are from the AMV collection; those below each line are from the HIV-1 collection. Base substitutions are indicated by a letter representing the new base found in the viral-template-strand RNA. For substitutions to a template T, the T is shown rather than a U, because these mutants were recovered as DNA, not RNA. For frameshift mutations, a solid triangle represents base addition, and an open triangle represents base loss. For additions or deletions within homopolymeric runs, it is not known which nucleotide within the run was added or deleted. One addition in the HIV-1 collection was of a cytidine between a uridine at position 147 and a guanosine at position 148. The position of the complementary oligonucleotide used to prime reverse transcription is indicated by the overlined DNA sequence. Polymerization is from right to left. For simplicity, only mutants having single-nucleotide changes are shown; the balance contained either more than a single change or changes in the positions where the primer hybridizes.

Table 2. Base substitution and frameshift error rates with DNA and RNA templates

Error	Position	DNA template* Occurrence	DNA template* Error rate, no. × 10^{-4}	RNA template† Occurrence	RNA template† Error rate, no. × 10^{-4}	Error rate ratio, DNA/RNA
Base substitutions						
T → C	−36	111	160	1	2.2	73
G → A	90	20	30	1	2.2	14
All other sites		55	0.7	18 (2)	0.3 (0.07)	2.3
One-nucleotide deletion						
−T (or U)	137–139	98	48	2	1.4	34
−C	106–108	46	23	0	≤0.7	≥33
−G	88–90	44	22	0	≤0.7	≥31
−T (or U)	70–73	19	7	3	1.6	4.4
−C	158	0	≤1.5	29 (0)	63 (≤4)	≤0.02
One-nucleotide addition						
+T (or U)	70–73	23	8	2 (0)	1.1 (≤1)	7.3
+A	91–94	5	1.8	6 (40)	3.2 (4)	0.6
+T (or U)	137–139	2	1.0	35 (12)	29 (17)	0.03

*DNA-templated mutation data are from ref. 4, and error rates were calculated as described. A total of 439 mutants was analyzed, and the average mutant frequency was 390 × 10^{-4}.

†Sequences of 139 randomly chosen, RNA-templated mutants from experiments 1 and 3 in Table 1 were determined, and error rates were calculated. For example, the base-substitution error rate of 3 × 10^{-5} was calculated by multiplying percentage of base-substitution mutants (18/139) by mutant frequency (150 × 10^{-4}, the average of experiments 1 and 3 from Table 1) and then dividing by the probability of expressing errors in the minus strand (50%) and by 114, the number of known detectable sites. Rates for errors at specific sites were calculated similarly, by using the data in Fig. 3 and four additional mutants (data not shown) that contained more than a single change. Numbers in parentheses are results obtained with AMV RT.

to distinguish between the other two sources, transcription and reverse transcription errors, we performed a parallel cDNA synthesis reaction with an AMV RT, followed by sequence analysis of mutants. As for HIV-1 RT, AMV RT catalyzed efficient cDNA synthesis (data not shown). The mutant frequency resulting from AMV RT synthesis with RNA was lower than for HIV-1 RT (Table 1). The number of occurrences and the rates for several types of errors are given in parentheses in Table 2, following the corresponding results with HIV-1 RT on RNA.

Our logic for comparing HIV-1 RT to AMV RT was that a higher rate for either RT over the other at any site or for any class of errors would imply that these were reverse transcription errors, whereas similarities in site-specific or class-specific error rates between HIV-1 and AMV RT would represent either reverse transcription errors common to both RTs or transcription errors present in the RNA template preparation. Based on this logic, four inferences on the origin of the mutants in the collection shown in Fig. 3 could be made. (*i*) The base-substitution error rate with RNA (Table 2) was 4.2-fold higher for HIV-1 RT (3 × 10^{-5}) than for AMV (7 × 10^{-6}). This result suggests that >80% of the base substitutions of Fig. 3 were generated by HIV-1 RT during cDNA synthesis. (*ii*) When comparing the two error rates for HIV-1 or AMV RT, the lower value provides a maximum estimate of the T7 RNA polymerase transcription error rate. (*iii*) The error rates were similar for HIV-1 and AMV RT (Table 2) for the two most common errors, addition of uridine at the UUU run or adenosine at the AAAA run (Fig. 3). Thus, either these mutants arose during transcription, which is known to generate pseudo-templated addition errors (13), or they reflect the common tendency of both HIV-1 and AMV RTs to commit addition errors at these homopolymeric template positions. The latter possibility is disfavored (but not excluded) by the observation that the two RTs have very different error rates for 1-nt additions with DNA templates. (*iv*) A mutational hot spot was also seen in the RNA collection at positions 158 and 159 (Fig. 3). Twenty-four mutations were found at these sites [Fig. 3 and mutants (data not shown) containing more than a single change], which are the first (position 159) and second (position 158) template nucleotides that would be encountered by the RT after initiating cDNA synthesis from the 3′-OH provided by the oligonucleotide primer. Five other mutants (data not shown) contained complicated frameshifts with one endpoint at these same nucleotides. Only one mutant in the AMV collection was observed at these positions (Fig. 3).

DISCUSSION

The results of the present study indicate that the overall fidelity of reverse transcription by HIV-RT is several-fold higher than is synthesis using a DNA template (Table 1). Just as with a DNA template (4), errors with RNA are nonrandomly distributed (Fig. 3). However, the location and types of mutants recovered with the two templates is different (Table 2), leading to differences in specific error rates.

We previously observed DNA-templated substitution and 1-nt deletion hot spots at homopolymeric sequences (4). Because the deletions likely result from template–primer slippage (4, 14, 15), the higher fidelity observed at these same sites with RNA (Table 2) suggests that misaligned intermediates are formed and/or used less frequently with an RNA template–DNA primer than with a DNA template–DNA primer. Based on the local nucleotide sequences and error specificities at the two DNA-templated substitution hot spots, we also suggested that some of these errors result from a dislocation mechanism, wherein a substitution is initiated by template–primer slippage (4, 16). The higher base-substitution fidelity at these hot spots (positions −36 and 90, Table 2) thus further suggests that misaligned intermediates are formed and/or used less frequently during synthesis with an RNA template–DNA primer than with a DNA template–DNA primer.

The hot spot for errors at the first and second positions 3′ to the 15-mer DNA primer (Fig. 3) is interesting. Its absence from the AMV collection suggests that these are HIV-1 RT errors rather than transcription errors. Thus, incorporation of the initial nucleotides by HIV-1 RT is less accurate than

subsequent polymerization, possibly reflecting aberrant enzyme–template–primer interactions. In studies using poly(rA)·oligo(dT), Reardon *et al.* (17) have suggested an abrupt transition in the rate constant for formation of the heterodimeric HIV-1 RT–template–primer complex as the primer length increases from 14 to 16 nt. Majumdar *et al.* (18) have demonstrated that, although polymerization by HIV-1 RT with poly(rA)·oligo(dT) is processive overall, the probability of template–primer dissociation was substantially greater after the first incorporation than after subsequent incorporations. We have previously reported a positive correlation between processivity and frameshift fidelity during DNA-dependent synthesis by HIV-1 RT (4). These observations suggest that the hot spot may reflect weak binding and/or lower processivity for the first few nucleotides incorporated.

The other hot spots with the RNA template are additions. These additions could have been generated by HIV-1 and AMV RT or they could reflect T7 RNA polymerase errors. The latter possibility is suggested by the fact that additions at homopolymeric runs are known to be produced during transcription (for review, see ref. 13). Here too, the relationship between fidelity and processivity is of interest (19).

The lowest mutant frequencies obtained after reverse transcription can be used to calculate the maximum transcription error rate. Under the conditions used here the substitution error rate per detectable nucleotide polymerized is $\leq 7 \times 10^{-6}$. The error rate for 1-nt frameshifts (135 known detectable sites), calculated from AMV RT data, is $\leq 9 \times 10^{-6}$ for 1-nt deletions, and $\leq 6 \times 10^{-5}$ for 1-nt additions. Excluding the two addition hot spots (Table 2) reduces the addition error rate to 1.3×10^{-5}. The strategy in Fig. 1 can be used in the future to further examine transcription fidelity and the parameters that influence it. Given the role of RNA polymerase II in the HIV-1 life cycle, it would be interesting to examine the fidelity of this polymerase.

Hubner *et al.* (10) have compared HIV-1 RT base-substitution fidelity in ϕX174 DNA versus RNA of the same sequence. In contrast to the results presented here, they conclude that synthesis with an RNA template is 7- to 20-fold less accurate than with a DNA template, at least for two mispairs that revert an amber codon. Ji and Loeb (9) have also developed a system to examine the fidelity of reverse transcription. From results obtained using a T3 RNA polymerase transcript of the same target sequence used in the present study, they suggest that the HIV-1 RT error rate for copying an 80-nt RNA template is comparable to that obtained in our earlier study with DNA (4). Possible explanations for the differences in the three studies include variations in experimental strategies, polymerization conditions, HIV-1 RTs used, and types and positions of errors that have been compared.

Can the RT error rates *in vitro* explain the extensive G → A hypermutation in the viral (+) strand that has been observed in cultured HIV-1 (20) and spleen necrosis virus (21)? Comparisons are limited by the possible influence of replication accessory proteins, phenotypic selection *in vivo*, and differences in the sequences and viruses being compared. However, if the G → A substitutions *in vivo* were replication errors with RNA, they would result from misincorporation of dTTP opposite template guanosine. The rate for this error in the present study is 2.8×10^{-5}. Generating G → A mutations during DNA-dependent replication in the life cycle would require misincorporation of dATP opposite cytidine. The rate for this error is 2.6×10^{-5} (4). These rates are >700-fold lower than the rate of 2×10^{-2} for G → A mutants per base pair per replication cycle seen with the spleen necrosis virus system (21). Although a rate for G → A substitutions per replication cycle is not yet available for HIV-1, Vartanian *et al.* (20) have suggested that the probability of finding a hypermutated HIV-1 genome is even greater than for the spleen necrosis virus. Thus, the error rates for the wild-type RT *in vitro* may be insufficient to explain the high rate and particular specificity (G → A changes primarily at GpA sites) of hypermutagenesis in HIV.

The differences in error rates with RNA and DNA templates suggest that minus-strand and plus-strand replication errors may not contribute equally to the final mutation rate and sequence diversity of the HIV and, perhaps, retroviruses in general. In a simple model of the retroviral life cycle, involving minus- and plus-strand synthesis followed by integration and transcription, plus-strand progeny would be templated from the product of minus-strand replication. In this instance, RNA template-directed errors would yield mutant progeny. Including all 20 substitution mutants from Table 2, the RNA template-dependent substitution error rate by HIV-1 RT in this study is no higher than 3.8×10^{-5}. If operative *in vivo*, this rate would account for 0.38 substitution error per 10 kb HIV-1 genome per replication cycle. This rate is >100-fold higher than the average mutation rate per genome in DNA-based microbes (22). However, after viral replication of both strands, cellular processes, such as replication or mismatch repair, could allow errors generated during DNA-dependent plus-strand synthesis to contribute to the mutant population. Including all 186 substitution mutants from Table 2, the average DNA template-dependent substitution error rate by HIV-1 RT is 2.4×10^{-4}. Thus, a small amount of information transfer from plus-strand synthesis products to progeny could influence the mutation rate.

We thank Dinh C. Nguyen for assistance with the sequence analysis and John W. Drake and William K. Kaufmann for critical evaluation of the manuscript. This work was supported by the National Institutes of Health AIDS Targeted Antiviral Program.

1. Saag, M. A., Hahn, B. H., Gibbons, J., Li, Y., Parks, E. S., Parks, W. P. & Shaw, G. M. (1988) *Nature (London)* **334**, 440–444.
2. Roberts, J. D., Bebenek, K. & Kunkel, T. A. (1988) *Science* **242**, 1171–1173.
3. Preston, B. D., Poiesz, B. J. & Loeb, L. A. (1988) *Science* **242**, 1168–1171.
4. Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H. & Kunkel, T. A. (1989) *J. Biol. Chem.* **264**, 16948–16956.
5. Weber, J. & Grosse, F. (1989) *Nucleic Acids Res.* **17**, 1379–1393.
6. Perrino, F. W., Preston, B. D., Sandell, L. L. & Loeb, L. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8343–8347.
7. Ricchetti, M. & Buc, H. (1990) *EMBO J.* **9**, 1583–1593.
8. Takeuchi, Y., Nagumo, T. & Hoshino, H. (1988) *J. Virol.* **62**, 3900–3902.
9. Ji, J. & Loeb, L. A. (1992) *Biochemistry* **31**, 954–958.
10. Hubner, A., Kruhoffer, M., Grosse, F. & Krauss, G. (1992) *J. Mol. Biol.* **223**, 595–600.
11. Kunkel, T. A. & Soni, A. (1988) *J. Biol. Chem.* **263**, 4450–4459.
12. Kunkel, T. A., Bebenek, K. & McClary, J. (1991) *Methods Enzymol.* **204**, 125–139.
13. Jacques, J.-P. & Kolakosky, D. (1991) *Genes Dev.* **5**, 707–713.
14. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. & Inouye, M. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 77–84.
15. Kunkel, T. A. (1990) *Biochemistry* **29**, 8003–8011.
16. Kunkel, T. A. (1985) *J. Biol. Chem.* **260**, 5587–5596.
17. Reardon, J. E., Furfine, E. S. & Cheng, N. (1991) *J. Biol. Chem.* **266**, 14128–14134.
18. Majumdar, C., Abbotts, J., Broder, S. & Wilson, S. H. (1988) *J. Biol. Chem.* **263**, 15657–15665.
19. Libby, R. T. & Gallant, J. A. (1991) *Mol. Microbiol.* **5**, 999–1004.
20. Vartanian, J.-R., Meyerhans, A., Asjo, B. & Wain-Hobson, S. (1991) *J. Virol.* **65**, 1779–1788.
21. Pathak, V. K. & Temin, H. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6019–6023.
22. Drake, J. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 7160–7164.